



NORTH-HOLLANDTechnological Forecasting & Social Change
70 (2002) 83–101

**Technological
Forecasting and
Social Change**

Disaggregative policy Delphi Using cluster analysis as a tool for systematic scenario formation

Petri Tapio^{*,1}

*Unit for Environmental Research in Finland Futures Research Centre (FFRC), Turku School of Economics
and Business Administration, PO-Box 110, Fin-20521 Turku, Finland*

Received 9 November 2000; received in revised form 3 July 2001; accepted 27 September 2001

Abstract

A critical phase of scenario making is the choosing of scenarios. In the worst case, a futures researcher creates scenarios according to his/her subjective views and cannot see the real quality of the study material. Oversimplification is a typical example of this kind of bias. In this study, an attempt towards a more data sensitive method was made using Finnish transport policy as an example. A disaggregative Delphi method as opposed to traditional consensual Delphi was applied. The article summarises eight Delphi pitfalls and gives an example how to avoid them. A two-rounded disaggregative Delphi was conducted, the panelists being representatives of interest groups in the traffic sector. Panelists were shown the past development of three correlating key variables in Finland in 1970–1996: GDP, road traffic volume and the carbon dioxide emissions from road traffic. The panelists were invited to give estimates of their organisation to the probable and the preferable futures of the key variables for 1997–2025. They were also asked to give qualitative and quantitative arguments of *why* and the policy instruments of *how* their image of the future would occur. The first round data were collected by a fairly open questionnaire and the second round data by a fairly structured interview. The responses of the quantitative three key variables were grouped in a disaggregative way by cluster analysis. The clusters were complemented with respective qualitative arguments in order to form wider scenarios. This offers a relevance to decision-making not afforded by a nonsystematic approach. Of course, there are some problems of cluster analysis

* Finland Futures Research Centre, PO-Box 110, Fin-20521 Turku, Finland. Tel.: +358-9-5613-6742; fax: +358-2-233-0755.

E-mail address: petri.tapio@tukkk.fi (P. Tapio).

¹ <http://engltpatio.htm>, <http://www.tukkk.fi/tutu>.

used in this way: The interviews revealed that quantitatively similar future images produced by the panelists occasionally had different kind of qualitative background theory. Also, cluster analysis cannot ultimately decide the number of scenarios, being a choice of the researcher. Cluster analysis makes the choice well argued, however.

© 2002 Elsevier Science Inc. All rights reserved.

1. Need for systematic scenario formation

One critical question concerning scenario formation is how to form scenarios relevant to decision-makers? How to make sure that the scenarios are not only futurist's prejudices or his/her own subjective ideas of the policy options? What are the relevant ranges of the values of the variables between scenarios?

The aim of this paper is to illustrate that one way to tackle this problem is to use the Delphi method applied in a disaggregative way. 'Disaggregative' here means that the goal of consensus is not adopted but the responses are grouped to several clusters by using cluster analysis as a systematic tool for grouping the core quantitative variables. The rather bare-boned clusters are then complemented with the respondent's qualitative arguments in order to construct more holistic scenarios. In this study, Delphi panelists are representatives of interest groups instead of individual professionals because the approach is considered an intermediate between expert poll and committee work in policy formation. The approach also raises some new problems, which are further discussed. The methodology is illustrated by the Delphi data gathered of the transport policy of Finland for 2025.

1.1. Case: Finnish transport policy

The disaggregative Delphi application of this study focuses on the future of the volume of economic output, traffic and environment. The problematique was reduced to three key variables:

- the gross domestic product measured in market exchange rates (GDP_{mer}) in real terms
- the road traffic volume measured in vehicle kilometers
- the carbon dioxide (CO_2) emissions of road traffic measured in metric tonnes

A decoupling of GDP and road traffic volume would imply immaterialisation or non-material economic growth in more traditional vocabulary. The decoupling of road traffic volume and the CO_2 emissions of road traffic would imply dematerialisation, or more simply, technical development of vehicles. In spite of all the writing about technical development, dematerialisation, decarbonisation and a qualitative change in the economy, the three key variables correlated strongly in 1970–1996 in Finland, especially from 1978 to 1996 [1–3] (Fig. 1).

The purpose of the application was to produce alternative scenarios focusing on these key variables for 1997–2025, which would be relevant from the point of view of the interest

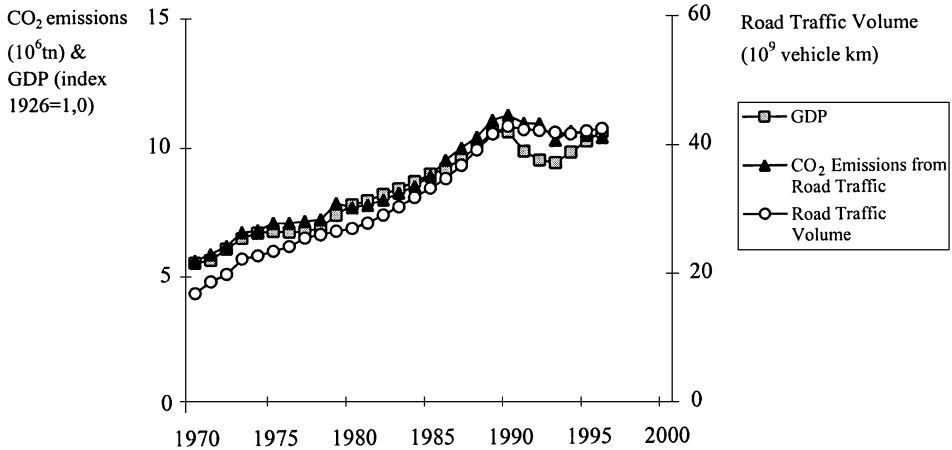


Fig. 1. The volume and CO₂ emissions of road traffic and the GDP_{mer} of Finland in 1970–1996.

groups of transport and environmental policy. The Delphi panel consisted of fourteen interest groups.

2. Avoiding Delphi pitfalls by disaggregation

To produce future scenarios for the economy, transport and environment an application of the Delphi method of two rounds was used. The idea of the traditional Delphi was to get an expert panel estimation of probable future on a topic that has many interpretations and is hard to formalise in mathematical models. It may also be used when there is insufficient data on the topic or when changes in the relations between variables are intuitively expected [4–9].

Especially the many interpretations and an expected change in the relations of the key variables are at present in the case of carbon dioxide policy of traffic sector. Programmes and discussions about it are conducted at many levels and organisations worldwide in UN [10], OECD Environmentally Sustainable Transport (EST) project [11], EU, US and Japan. In connection to EU traffic policy, the Ministry of Transport and Communications of Finland also organised a Working Group of CO₂ Emissions [12].

The Delphi method is an iterative process consisting of at least two rounds (sometimes even five). The purpose of the multiple rounds is to give panelists feedback from the previous rounds where experts make arguments and/or evaluations for some issues to happen. The argumentative rounds are usually anonymous, so that the status or background organisation of the experts would not affect others' opinions. The ideal of the procedure is that the panelists would make some tacit knowledge explicit and that the best argument should win. Another ideal of the traditional Delphi was to reach a consensus among the panelists [4,7,9].

As the readers of this journal probably remember, the traditional Delphi method was subjected to severe criticism in the mid-1970s. Several alternatives to Delphi were promoted, such as Shang inquiry by Ford [13], POSTURE by Brockhaus [14] and SPRITE by Bedford

[15]. Also, Hill and Fowles [16] wrote critically about Delphi in the special issue of *Technological Foresight and Social Change* (TFSC). Even the pro-Delphi scholars like Turoff [17] and Linstone [18] were concerned about the many pitfalls of the traditional Delphi.

Another wave of criticism can be found from the pages of TFSC in the late 1980s and early 1990s. Benarie [19] discussed the philosophical limitations of Delphi and Delphi-like methods. Woudenberg [20] and Kastein et al. [21] criticised poor reliability and accuracy of Delphi forecasts in comparison to other methods. Rowe et al. [22] observed inadequacies in the actual questions and feedback of Delphi applications. Webler et al. [23] were especially worried about panelists' lack of commitment to the process and developed Delphi towards a committee work. As the first wave of Delphi critique more or less totally questioned the worth of Delphi, the second wave merely promoted modifications for better study design.

However, Delphi survived the two waves of criticism as can be seen in the extensive bibliography of Gupta and Clarke [24]. The following text summarizes eight Delphi pitfalls from the critical discussion and explains how the pitfalls were dealt with in this study.

2.1. Biased selection of the panelists

Delphi critique often remarks that in applications little effort is put to a reliable selection of the panelists [15,16,18]. Often used conomination tends to result in a biased sample, because experts apparently conominate colleagues that represent similar school of thought. Cuhls [25] has suggested that conomination is a good start, but certain basic background factors, such as sex, age and professional background, should be checked before the Delphi manager can safely stop looking for new panelists. She also suggests scanning of publications, institutions and public databases relevant on the study object to get reliable samples. For example, the panel of a recent Delphi study "The Future of Mobility" consisted of 96% of men, only 14% of under 40-year-old experts and 51% of technically or economically educated experts [26].

In this study, the panelists were selected in a somewhat unusual way. Instead of individual experts, the participants were fourteen interest groups that have an interest to the transport and environmental policy in Finland. The rationale behind this choice is that the disaggregative Delphi serves as an alternative to committee work in forming normative policies but aims not at replacing policy makers with scientists. The organisations represented the following categories:

- Traffic administration
 - Ministry of Transport and Communications
 - Rail Administration
 - Road Administration
- Environmental administration
 - Ministry of the Environment
- Local administration
 - Helsinki Metropolitan Area Council/Transportation Department (manager's own view)
- Lobbying groups of different traffic modes
 - Bus Transport Federation (bus transport)

- Automobile and Touring Club of Finland (passenger cars)
- Traffic League (surface public transport and soft modes)
- Traffic Policy Association Majority (soft modes)
- Other groups with economic interest
 - The Confederation of Finnish Industry and Employers (interest in freight transport)
 - Finnish Road Association (road construction)
 - Transport Workers' Federation (trade union, e.g., road haulage drivers and bus drivers)
- Environmental group
 - Dodo—The Living Nature of the Future (a group of young students)

Ministry of Treasury, the car import organisation and an older environmental group Finnish Nature Conservation Federation dropped out in the first round because of lack of time and/or feeling that their views would be presented by some of the other panelists. Water and air transport related interest groups were not involved, because for geographical reasons they do not compete much with road traffic in Finland. The lack of an air transport interest group can be considered a bias, however.

2.2. *Disregarding organisations*

Anonymity is usually maintained among Delphi panelists in order to bring out more honest views without having to be afraid to lose face or a job. On the other hand, individuality and anonymity have been claimed to be reasons for the lack of commitment consequently resulting in high dropout rates, scarce written arguments and hasty “snap-judgments” instead of cautious ponder and analysis of the issue [15,23]. One solution worth a try could be to ask panelists to act as representatives of their organisations instead of as individual experts. The author has not seen references suggesting this. In this study, organisational representation was experimented with, mainly because the application was designed to improve the process of typical committee work where participants do represent their organisations.

The representatives of the organisations were sampled systematically by making a phone call to the operational top managers of the organisations. After that, the organisation was free to work in its own way to appoint a representative or representatives. Some managers participated themselves whereas some delegated the task to their subordinates, who were phone-called as well. Three organisations appointed two representatives and some of the other respondents may have asked their colleagues or bosses for second opinion. The representatives of three other organisations changed between the rounds, one because of lack of personal familiarity to quantitative approach, one because of retirement and one because of change of employer. The change of opinion of these three organisations between the rounds was not different from the other organisations. The respondents were asked for their organisation's view on the most probable and the most preferable future. This is in line with the Policy Delphi applications [7,17,27].

Focusing on the organisations instead of individuals brought out interesting features. Some representatives protested that their organisation did not have official quantitative statements on the carbon dioxide emissions of road traffic or GDP. They were explained that no official declaration was asked for, merely well argued estimates and evaluations. This sufficed for most participants but two managers admitted only to represent their own views instead of the organisation's. Some respondents complained that their own view differs from what they regarded as the organisation's view. In these cases, the organisation's view was asked for anyway.

It was not the focus of the study to analyse the average opinion of an organisation. It might differ significantly from the organisation's opinion, because organisation is something else than just a sum of its individuals. There seemed to be discrimination in the selected organisations, since only 2 women as opposed to 12 men were gathered by this method.

2.3. Forgetting disagreements

Applications of Delphi have been criticised of ignoring and not exploring disagreements, which could generate artificial consensus [4,15–17]. In this study, a goal of consensus between the participants was not adopted. Instead, a set of alternative long-term traffic and environmental policy scenarios were produced. This type of unconsensual, or disaggregative, Delphi has been applied already in the 1970s [28,29]. Also, consensus seemed to be rather unimportant goal in the big national technology foresight studies conducted by Delphi in the 1990s [30] (the conclusion is less clear in Ref. [31]). Arguments for a more disaggregative Delphi have been stated by other Delphi users as well [27,32]. However, most of the articles of the Delphi book of Adler and Ziglio [33] still seemed to have a consensus approach.

2.4. Ambiguous questionnaires

Delphi literature is full of warnings and descriptions of poorly formed questionnaires (e.g., Refs. [7,9,15,22,34]). The actual questions have been especially criticised for being too abstract and leaving room for too many interpretations. In the first Delphi round the study material was gathered with a questionnaire that included a figure of the development of GDP, road traffic volume and CO₂ emissions from road traffic in 1970–1996 in Finland. The panelists were asked to manually draw the trends for the most probable future and the most preferable future for 1997–2025. The questionnaire was pretested with and commented by four traffic professionals (see Ref. [34]).

2.5. Oversimplified structured inquiry

Another pitfall in Delphi applications is oversimplified structured inquiry that does not leave room for new ideas [16,18,30]. To avoid this, the panelists were asked to write down arguments in an open form why and how they expected the most preferable future to happen

and why they believed the most probable future would happen. The most preferable future was defined as a future that is technically possible and desirable taken into account all the aspects and impacts relevant for the organisational viewpoint. The definition was a tryout to help the panelists separate themselves from too conservative or pessimistic views (“daily realism”) concerning social change without sliding to total science fiction. The definition was unfortunately ambiguous, one interpretation covering only direct utility of the organisation and another enabling a more altruistic approach as well. Another requirement was that the answers should form a coherent scenario instead of isolating the three variables from each other or from their economic, social and ecological context.

2.6. *Feedback reports without analysis*

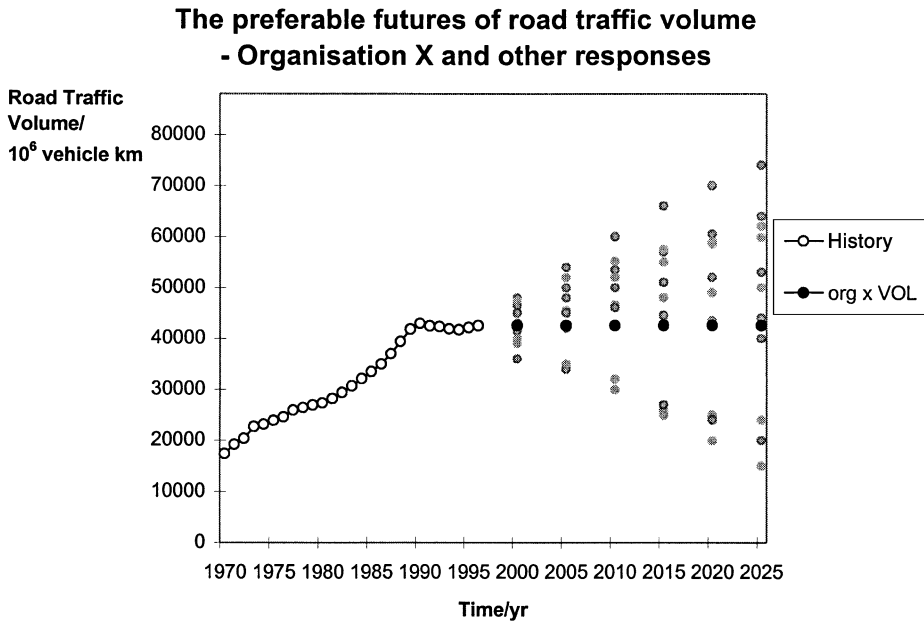
Scarce feedback and summary reports are another criticised feature of Delphi studies [15,18,30,34]. Thus, a summary report was produced that presented all the responses of the three variables separately and included the arguments produced by the open form. Every respondent could see his/her own organisation's first round answer emphasised with a different colour in relation to other responses, which is illustrated in Fig. 2.

2.7. *Forgetting the arguments*

Although the fundamental ideal of Delphi is that the best argument should win, in actual applications arguments have not had a central role [17,30,35]. To get more in-depth arguments, the second Delphi round in this study was conducted by interviews. Because the open questionnaire of the first round produced less material than was hoped, a more structured thematic interview was conducted in the second round.² The interviewee was systematically asked to comment on the arguments presented by estimations of lower and upper curves than their own responses. The respondents were encouraged to give arguments supporting their view and allowed to change their answers from the first round on the basis of the arguments of the other respondents. This is a similar feature to the “Argument Delphi” developed by Kuusi [30], originating partially from argumentation rules of van Eemeren et al. [38].

The role of the Delphi moderator was to state contra-arguments in order to get more in-depth arguments from the respondents. This approach immediately leads to doubts of bias originating from the explicit and implicit ways in which the contra-arguments are presented by the interviewer. Four ways to ameliorate this effect were employed: (a) in the beginning of the interviews, it was made clear that the method includes presenting other panelists' contra-arguments; (b) the interviewer tried to isolate himself from the arguments by expressions such as “often a contra-argument is stated to the point you are making that...”; (c) in the earlier interviews, the interviewer presented also hypothetical contra-arguments on the basis of literature and 8 years experience in the transport field;

² Thematic interviews originate from the “focused interview” of Merton and Kendall [36]. Merton himself would have preferred the concept “focussed interview” [37].



Arguments for the upper curves

- public transport should be supported without restricting passenger car traffic
- the freedom for private car use should not be restricted
- the growth of CO₂ emissions can be stopped with technical development

Arguments for the middle curves

- the growth potential of passenger car traffic should be guided towards soft modes, public transport and telecommunications
- urban infill is preferable to reduce the demand for traffic
- traffic is a mean, not an end itself, economy and communication should be handled with low need for traffic
- economic growth should be achieved by electronic industry and services, which would reduce the growth rate of freight transport in relation to GDP

Arguments for the lower curves

- telecommunications will and should substitute physical traffic
- local more non-material economy is preferable
- railroads should be emphasised in both passenger traffic and freight transport
- CO₂ emissions should be reduced 60-80% from today's level to stop climate change, which is not possible without also reducing road traffic volume

Fig. 2. Illustration of round one results to a particular organisation x concerning preferable estimates of road traffic volume of Finland. (The apparent erroneous way to present the volume scale with five numbers originates from the first round questionnaire and was deliberately repeated in the second round, because the choice of scale might have had influence on the second round responses.)

(d) in the later interviews, the cumulative total chain of arguments and contra-arguments presented to a certain issue was dealt with. The interviews lasted from 1.25 to 4 h and the total tape-recorded material was approximately 35 h.

The methodological concern to make the interviews rational well-argued discussions seemed to be successful. A lot in-depth arguments were produced, such as reference to specific research or statistics and revelation of the social theory of mobility behind the answer, etc. These are difficult to get with a questionnaire or an interview without contra-arguments. The interview strategy adopted in this study has features of ethnographic decision-tree modeling, although here decisions were not computerised to a yes/no dichotomy (see Ref. [39]).

As a rule, exploratory thematic interviews should be made first and then, based from the interview, form a more exact questionnaire for the next round(s).³ This would clarify concepts, improve the relevance of the questionnaire and increase motivation to participate. There were three reasons to break the rule in this study. First, the starting point of the study, i.e., the three curves in Fig. 1, was so illustrative, that respondents were (correctly) expected to be motivated to answer. Second, the respondents had a high motivation because they worked directly with the issues. Third, why conduct argumentative interviews if there are no process borne arguments nor statements to argue about?

2.8. *Lack of theory*

Delphi users have often been criticised for the lack of theoretical understanding of the methodological procedure and less often for the lack of theoretical framing of the substance [16,35,41]. The harshest conclusion has probably been stated by Bell [41]:

So far, Delphi researchers use, create, test or know precious little—if any—social theory. (p. 270)

This critique seems overly harsh. Especially the philosophical and theoretical foundations of the Delphi procedure have been considered quite often. To mention a few Scheele [42] made an effort to place Delphi in the context of phenomenological epistemology; Mitroff and Turoff [43] presented five philosophical inquiry systems and stated that Delphi has a role to play in all of them, but the applications might be different from each other; Rowe et al. [22] placed Delphi in the context of judgment and decision-making theories (in business administration); Kuusi [30] developed a “general theory of consistency” as a philosophical framework and analysed how Grupp has tested certain theories of technological paradigms. There are also obvious connections of Delphi to Habermasian “undistorted communication” and critical pragmatist planning theory, but they are impossible to analyse in detail here (see, e.g., Refs. [44,45]).

Actual testing of social theories with Delphi applications is not as common. In this study, the responses were interpreted in the light of theories of environmental policy, but the analysis will be made in a separate paper. Although the critique of lack of theory is overestimated, it still is a relevant pitfall to keep in mind.

³ The rule is not recommended by Hirsjärvi and Hurme, however [40]. They state that the research problem is more important than any methodological principle and present five equal approaches: (1) first qualitative then quantitative; (2) first quantitative then qualitative; (3) iterative process; (4) both methods at the same time but with different respondents; (5) both methods at the same time with one respondent.

Table 1

Comparison of traditional Delphi and disaggregative Delphi and the extra characteristics of this study

Feature	Traditional Delphi	Disaggregative Delphi
<i>Similarities</i>		
Ideal		The best argument wins
Transparency		Anonymity of arguments
Iterativity		Multiple rounds
<i>Differences</i>		
Philosophy	Consensus	Dissensus
Goal	Accurate prediction	Alternative scenarios
Feedback	Median and interquartiles	All responses and key arguments
Statistical test	e.g., analysis of variance	e.g., cluster analysis
<i>Extra characteristics of this study</i>		
Case	Individual professional	Interest group
Transparency after the study	Anonymity is retained also after the Delphi rounds	Anonymity is limited to the argument phase
Form of data	Questionnaire	Questionnaire and interview

The features of the traditional Delphi and the disaggregative Delphi are summarised in Table 1. The table includes also some extra features of this study that are different from the traditional Delphi.

3. Cluster analysis as a disaggregative tool

3.1. Critical choices in cluster analysis

The answers of the three key variables, GDP, road traffic volume and the CO₂ emissions from road traffic, were grouped in a disaggregative way. The grouping was made by cluster analysis provided in the SPSS 8.0 software. The idea of using cluster analysis in Delphi studies has been presented at least by Turoff and Hiltz [27]. Clustering methods do not require random sampling unless they are used in verifying a theory, because they can be used only as tools to group similar cases together [46,47].

The choice of the hierarchical clustering method is an exhaustively discussed topic in the literature of classification. The deeper one gets into this discussion, the less agreeable criteria one gets. There has been a number of validation studies on artificial data sets, where the true number of clusters is known. Then the different methods have been used and the recovery percentage of the methods are compared [48]. The only obvious agreement seems to be, that the nearest neighbour (i.e., single linkage) method should not be used unless the clusters are supposedly of chain shape [48,49].

Usually, the test data have been extensive, which is not the case in this research (only 24 cases because two respondents did not produce quantitative responses of all the three variables). A “cluster” here consists of only few cases and cluster size probably varies

which makes the choice of the more statistically oriented Ward's method questionable, although it has done well in comparisons mentioned above [50].

Henrion et al. [49] go as far as suggesting to run the analysis with all available clustering methods and pick up the one that makes sense. From the standard Popperian deductionist point of view adopted in statistical science, this would be circular reasoning [51]. But from the more Newtonian inductionist point of view adopted in many technical and qualitatively oriented social sciences, experimenting different methods would also seem as a relevant strategy, because of the Dewian “learning by doing” background philosophy (see Ref. [52]).

From this basis, and taken into account that the purpose of the clustering in this study is just to group a wide range of 24 responses together, the furthest neighbour (i.e., complete linkage) clustering method was chosen. Another consistent choice would have been the between groups (i.e., average linkage) method, which was tried for comparison to the furthest neighbour method with the first Delphi round data. There were no significant differences between the methods with the data.

The critical decisions in clustering include choosing clustering elements and variables [48]. The elements to be clustered in this case were the responses from the different respondent organisations. Both the preferable and probable responses were used simultaneously. The data were organised as eighteen variables to the SPSS program so that the values for GDP, road traffic volume and CO₂ emissions were dotted every 5 years in 2000–2025. The first variable was the GDP in the year 2000, the second variable the GDP in the year 2005 and the sixth variable the GDP in the year 2025. The seventh variable was the CO₂ emissions in 2000, the eighth the CO₂ emissions in 2005, etc., so that the 18th variable was the road traffic volume in 2025. The data is thus a time series of the three key variables. Clustering was made on the basis of these values and cluster centers illustrated the cores of scenarios. No other data than the three key variables of GDP, road traffic volume and CO₂ emissions of road traffic were used to form the clusters.

Besides the decision of clustering entities, the choice of a measure of association is essential [48,53]. The simple Euclidean distance was chosen because the variables are on a relative scale. The cases have reasonably similar shape of the different variables, which also supports the choice. No variable standardisation was conducted nor weights put on different variables. This emphasises the differences and similarities of road traffic volume because the volume numbers are greater than the CO₂ and GDP numbers. The choice is consistent with the emphasis of traffic in the scenarios.

3.2. Hierarchical cluster analysis with the case data

Next, I will present illustrative results of the respondents' views about preferable futures of GDP, road traffic volume and the CO₂ emissions from road traffic in Finland for 1997–2025, according to the second Delphi round. The quantifiable responses are grouped with the help of hierarchical cluster analysis in order to make the critical decision of the number of the clusters. Final cluster centres are illustrated and interpreted as the core of alternative traffic scenarios. The responses of the preferable and the probable futures are grouped simulta-

neously in order to find out the widest scope of views. The status of the responses (preferable or probable) grouped to a given cluster is shown in Fig. 3.

The dendrogram of the furthest neighbour clustering (Fig. 3) does not illustrate clearly enough where cases three and 19 are grouped. The vertical “icicle” in Table 2 manages to bring out the grouping better.

It is easy to agree with the general idea that the purpose of analysis is to simplify enough but not too much. The problem is to define what is enough and what is too much. In this study, the purpose is to simplify as little as possible in order to keep a wide range of policy alternatives inside the Delphi process as stated above.

Futures research literature often recommends that two alternatives should not be formed, because it implies a preference of one alternative over another. Also, there are limits to human

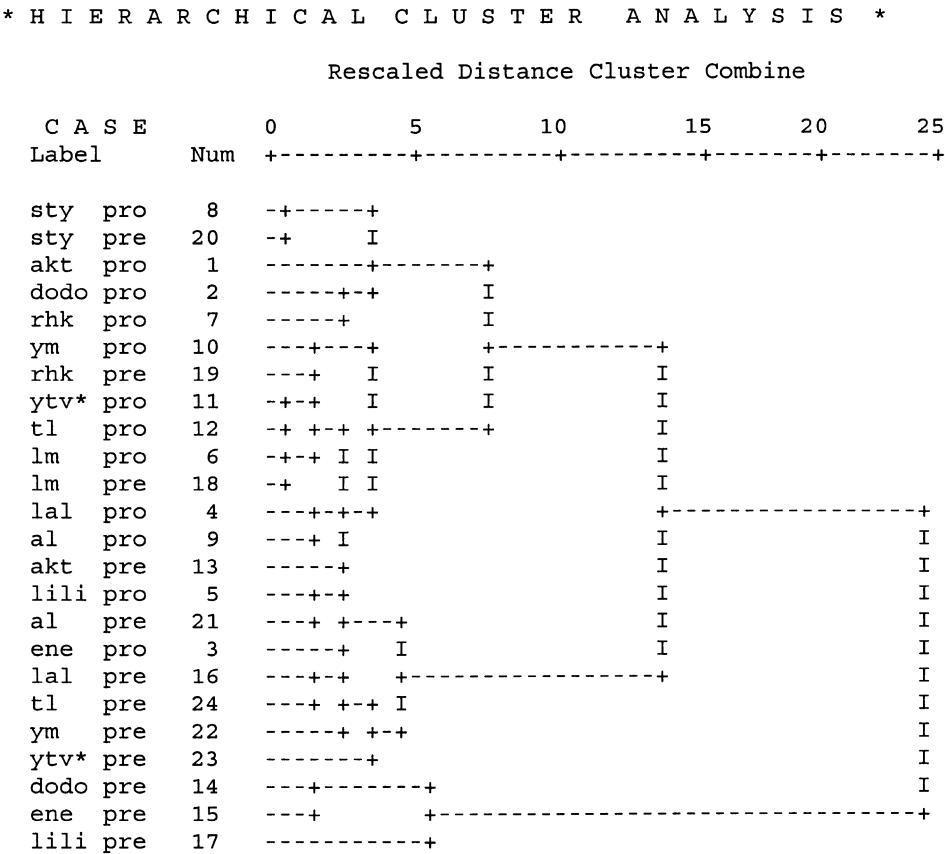


Fig. 3. Dendrogram using furthest neighbour (complete linkage) method labels: pro=probable, pre=preferable. STY=Finnish Road Association, AKT=Transport Workers' Federation, DODO=Dodo—The Living Nature of the Future, RHK=Finnish Rail Administration, YM=Ministry of the Environment, YTY*=Helsinki Metropolitan Area Council/Transportation Department (*manager's own view), TL=Finnish Road Administration, LM=Ministry of Transport and Communications, LAL=Bus Transport Federation, AL=Automobile and Touring Club of Finland, LILI=Traffic League, ENE=Traffic Policy Association Majority.

Table 2

The process of clustering the cases as vertical icicle

Number of clusters	Case																							
	17	15	14	23	22	24	16	21	5	3	19	10	13	12	11	18	6	9	4	7	2	20	8	1
1	xx	xx	xx	xx	xx	xx	xx	xx	xx	xx	xx	xx	xx	xx	xx	xx	xx	xx	xx	xx	xx	xx	xx	x
2	xx	xx	x	xx	xx	xx	xx	xx	xx	xx	xx	xx	xx	xx	xx	xx	xx	xx	xx	xx	xx	xx	xx	x
3	xx	xx	x	xx	xx	xx	xx	xx	xx	x	xx	xx	xx	xx	xx	xx	xx	xx	xx	xx	xx	xx	xx	x
4	xx	xx	x	xx	xx	xx	xx	xx	xx	x	xx	xx	xx	xx	xx	xx	xx	xx	xx	x	xx	xx	xx	x
5	x	xx	x	xx	xx	xx	xx	xx	xx	x	xx	xx	xx	xx	xx	xx	xx	xx	xx	x	xx	xx	xx	x
6	x	xx	x	xx	xx	xx	x	xx	xx	x	xx	xx	xx	xx	xx	xx	xx	xx	xx	x	xx	xx	xx	x
7	x	xx	x	x	xx	xx	x	xx	xx	x	xx	xx	xx	xx	xx	xx	xx	xx	xx	x	xx	xx	xx	x
8	x	xx	x	x	xx	xx	x	xx	xx	x	xx	xx	xx	xx	xx	xx	xx	xx	xx	x	xx	x	xx	x
9	x	xx	x	x	xx	xx	x	xx	xx	x	xx	xx	xx	xx	xx	xx	xx	xx	xx	x	xx	x	xx	x
10	x	xx	x	x	xx	xx	x	xx	xx	x	xx	x	xx	xx	xx	xx	xx	xx	xx	x	xx	x	xx	x
11	x	xx	x	x	xx	xx	x	xx	xx	x	xx	x	x	xx	xx	xx	xx	xx	xx	x	xx	x	xx	x
12	x	xx	x	x	xx	xx	x	xx	x	x	xx	x	x	xx	xx	xx	xx	xx	xx	x	xx	x	xx	x
13	x	xx	x	x	xx	xx	x	xx	x	x	xx	x	x	xx	xx	xx	xx	xx	xx	x	x	x	xx	x
14	x	xx	x	x	x	xx	x	xx	x	x	xx	x	x	xx	xx	xx	xx	xx	xx	x	x	x	xx	x
15	x	xx	x	x	x	xx	x	xx	x	x	xx	x	x	xx	xx	xx	x	xx	x	x	x	xx	x	x
16	x	x	x	x	x	xx	x	xx	x	x	xx	x	x	xx	xx	xx	x	xx	x	x	x	xx	x	x
17	x	x	x	x	x	xx	x	xx	x	x	x	x	x	xx	xx	xx	x	xx	x	x	x	xx	x	x
18	x	x	x	x	x	xx	x	xx	x	x	x	x	x	xx	xx	xx	x	x	x	x	x	xx	x	x
19	x	x	x	x	x	xx	x	x	x	x	x	x	x	xx	xx	xx	x	x	x	x	x	xx	x	x
20	x	x	x	x	x	x	x	x	x	x	x	x	x	xx	xx	xx	x	x	x	x	x	xx	x	x
21	x	x	x	x	x	x	x	x	x	x	x	x	x	xx	x	xx	x	x	x	x	x	xx	x	x
22	x	x	x	x	x	x	x	x	x	x	x	x	x	xx	x	x	x	x	x	x	x	xx	x	x
23	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	xx	x	x

capacity to outline alternatives; often seven alternatives is regarded as the maximum relevant number of scenarios. Schwartz et al. [28] state that three to six scenarios is most common. Mannermaa [9] suggests three to five scenarios. However, three is problematic since the decision-maker could easily prefer the middle one.

More recently, major scenarios have been constructed on differences of some key parameters and further divided into subscenarios based on less important parameters, which is an illustrative way to increase the capacity to handle alternatives (e.g., Refs. [54,55]). Also, scenarios can be built by making scenarios of factors external to decision-making and then cross-matrixing them to the scenarios of the object system, which are more decision-making oriented (e.g., Refs. [56–58]). This kind of hierarchisation has not been made in this study because the euclidean association measure and scale of variables already emphasise road traffic volume as the most important variable and because all the three key variables are dependent on decision-making.

In addition, external substantial theoretical categories could be used in deciding the choice of the number of clusters [46]. In the theoretical part of this research, five scenarios were produced [59], and it is sensible to make a comparison between clusters found in cluster

analysis and the theoretical scenarios. The comparison is made in another article that is more result-oriented [60].

On the other hand, if the sampling of different interest groups is representative, the clustering method is supposed to recover the real structure of the transport and environmental policy views. This possibility should be respected regardless of the expected behaviour of decision-makers or theoretical categories. This is a difficult dilemma for which it is hard to find reasonable solutions.

The dendrogram in Fig. 2 gives a well argued choice between three, four, six and ten clusters, four and six being the most apparent ones. Also, the choice of five clusters is possible but it would give a rather imbalanced size of the clusters. When comparing this information with the wide range of original responses, maybe cluster numbers of 3, 5 and 10 should be excluded. We are left with the option of four or six clusters, six giving a wider range of the clusters.

Derived from these criteria, the most relevant number of clusters seems to be six.

3.3. Final cluster centres

Now that similar cases have been grouped together, final cluster centres have to be calculated. They are simply the arithmetic means of each variable in different clusters produced by the hierarchical method. This is consistent with the choice of nonweighed Euclidean association measure.

On the basis of the hierarchical clustering six clusters are constructed. The Clusters 1, 2 and 4 are illustrated in Fig. 4 and Clusters 3, 5 and 6 in Fig. 5. Of course, the reader may think whether it would be sensible to further group Clusters 5 and 6 and on the other hand Clusters 3 and 4 to reduce the total number of clusters to four.⁴

Cluster 1 is a material-growth oriented slightly improved business as usual development, Cluster 2 is close to the characteristics of ecological modernisation, Cluster 3 presents the idea of structural change in economy with a rather modest growth rate, Cluster 4 also presents structural change but with a higher 'factor four' nonmaterial growth rate, Cluster 6 promotes a zero-growth steady-state economy and Cluster 5 even negative economic growth following the principles of the deep ecology school of thought [60].

3.4. Importing qualitative data to the clusters

The quantitative clusters of the key variables do not tell anything about the reasons why and how these developments would or should take place from the point of view of the respondents. The qualitative arguments presented in the interviews of the second Delphi round provide this data and can be reported in connection with the clusters.

A complete analysis of the qualitative data is not called for here (it is reported in Ref. [60]), but a few methodological considerations can be posed. An interesting question is whether

⁴ In fact, doing this would have produced clusters that would have been closer to the theoretical scenarios presented in the theoretical part of the research [59], but to me Clusters 3 and 4 as well as Clusters 5 and 6 really seem different. Without applying cluster analysis, this distinction would possibly never have occurred to me.

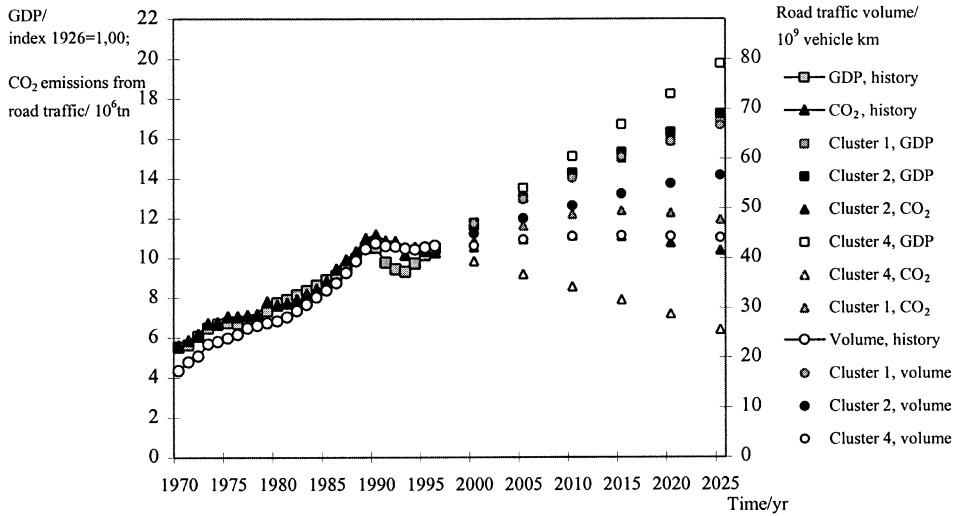


Fig. 4. The development of GDP, road traffic volume and the CO₂ emissions from road traffic 1970–1996 in Finland and Clusters 1, 2 and 4 for 1997–2025. (The squares stand for GDP, circles for road traffic volume and triangles for the CO₂ emissions from road traffic. The different clusters for future are illustrated with separate colours: Cluster 1: grey, Cluster 2: black and Cluster 4: white. The values of GDP and road traffic volume in Cluster 1 match exactly and cannot therefore be clearly seen in the figure.)

qualitative arguments of the responses within a cluster are similar? The answer is mostly yes but not necessarily.

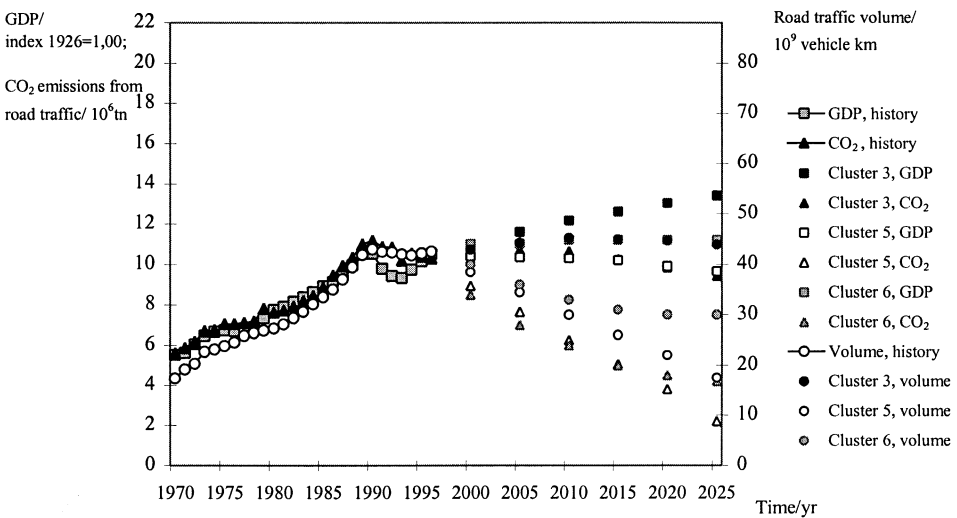


Fig. 5. The development of GDP, road traffic volume and the CO₂ emissions from road traffic in 1970–1996 in Finland and Clusters 3, 5 and 6 for 1997–2025. (The squares stand for GDP, circles for road traffic volume and triangles for the CO₂ emissions from road traffic. The different clusters for future are illustrated with separate colours: Cluster 6: grey, Cluster 3: black and Cluster 5: white.)

Certain factors were presented repeatedly, such as the success (or nonsuccess) of electronic, metal and forest industry, globalisation of economy, competitiveness of Finnish economy, areal and urban structure, transportation infrastructure, logistics, telecommunications, lifestyles and values, technical development and taxation policies of the state. But responses within the same cluster sometimes emphasise the factors very differently. For example, Cluster 2 included responses stating that the urban sprawl will continue as well as responses expecting urban infill. This makes the formation of a coherent scenario problematic.

Another problematic feature is that the structure of the responses of probable future seemed to differ from the responses of preferable future. The preferable future included more policy measures whereas the probable included more external factors to decision-making. Maybe the respondents regarded business as usual policy as probable and a more active policy as preferable. But this leaves open the question, why was there such a wide variety of responses concerning probable future as well.

4. Discussion

This paper described a systematic way to produce scenarios that would be relevant to decision-making. The methodology is based on a disaggregative Delphi application that uses cluster analysis as a tool for disaggregative grouping of the responses of the quantitative key variables from the Delphi panel. The quantitative statements were connected to the qualitative arguments in interviews of the second Delphi round. In sum, it seems that:

- It is valuable to gather both qualitative and quantitative data of the statements of the respondents and seriously analyse them in connection with each other.
- The argumentative role of the Delphi moderator brings up arguments that would else stay implicit.
- Cluster analysis contributes novel groupings of the responses that the researcher might not think of before the analysis (for example, the preferable future statements of Finnish Road Administration were clustered together with the statements of Ministry of the Environment instead of the Ministry of Transportation and Communication).
- Cluster analysis works well in the grouping and the qualitative interviews produce insight to the quantitative statements, but this does not automatically produce coherent scenarios which require extra heuristic work.
- However, the method presented here provides a good tool for the critical phase of scenario building, where the key characteristics of the scenarios are formed.

Acknowledgements

The author would like to thank Tarja Meristö for help in the study design, Kerstin Cuhls, Matti Kamppinen and Osmo Kuusi for helpful critical comments to an earlier version of the paper and Pekka Korhonen for guidance in cluster analysis. The second Delphi round was funded by the

Ministry of Transport and Communication and the Academy of Finland. The Academy of Finland and the Henry Ford Foundation funded participation to the Young Scientists' Summer Programme at the International Institute for Applied Systems Analysis (IIASA), where the author received a number of helpful comments for the design of the second round.

References

- [1] K. Mäkelä, Suomen liikenteen hiilidioksidipäästöt. The website of Technical Research Centre Finland, <http://www.vtt.fi/yki/yki6/liisa/co2dats.htm>, 1997. [The Carbon Dioxide Emissions from Traffic in Finland.]
- [2] R.A. Finn, Koko maan liikennesuorite autolajeittain vuosina 1970–1997/2020E. The website of Finnish Road Administration. Available at: <http://www.tieh.fi/aikas/liiks.htm>, 1997. [Road traffic volume by mode in the whole country in 1970–1997/2020 (forecast).]
- [3] Statistics Finland: Volyyymi-ideksit 1860–1995. The website of Statistics Finland, <http://www.stat.fi/tk/to/sarjat.html>, 1997. [Volume indexes 1860–1995 (of GDP of Finland).]
- [4] H.A. Linstone, M. Turoff (Eds.), *The Delphi Method. Techniques and Applications*, Addison-Wesley Publishing, Don Mills, 1975 (620 pp.).
- [5] W.E. Riggs, The Delphi technique. An experimental evaluation, *Technol. Forecast. Soc. Change* 23 (1983) 89–94.
- [6] G. Rowe, G. Wright, F. Bolger, Delphi: A reevaluation of research and theory, *Technol. Forecast. Soc. Change* 39 (1991) 235–251.
- [7] E. Ziglio, The Delphi method and its contribution to decision-making, in: M. Adler, E. Ziglio (Eds.), *Gazing into the Oracle. The Delphi Method and its Application to Social Policy and Public Health*, Jessica Kingsley Publishers, London, 1996, pp. 3–33.
- [8] A. Rotondi, D. Gustafson, Theoretical, methodological and practical issues arising out of the Delphi method, in: M. Adler, E. Ziglio (Eds.), *Gazing into the Oracle. The Delphi Method and its Application to Social Policy and Public Health*, Jessica Kingsley Publishers, London, 1996, pp. 34–55.
- [9] M. Mannerman, Tulevaisuuden hallinta-skenaariot strategiatyöskentelyssä, *Ekonomia-sarja*, Suomen Ekonomiliitto and WSOY, Porvoo, 1999, 227 pp. [Management of future—Scenarios in strategic management, in Finnish.]
- [10] IPCC: Climate change 1995. Impacts, adaptations and mitigation of climate change: Scientific–technical analyses, in: R.T. Watson, M.C. Zinyowera, R.H. Moss, D.J. Dokken (Eds.), *Contribution of Working Group II to the Second Assessment Report of the Intergovernmental Panel on Climate Change*, Cambridge Univ. Press, New York, 1996 (879 pp.).
- [11] OECD: Environmentally sustainable transport. Final Report on Phase II of the OECD EST Project. Volume 1 Synthesis Report. Environment Directorate, Environment Policy Committee, Working Party of Pollution Prevention and Control, Working Group on Transport, ENV/EPOC/PPC/T(97)1/ FINAL, 81789, available at <http://www.oecd.org/est>, 1997.
- [12] Ministry of Transport and Communications: Toimenpiteet tieliikenteen hiilidioksidipäästöjen vähentämiseksi, Publications of the Ministry of Transport and Communications 16/99, Edita, Helsinki, 1999, 80 pp. [Measures to decrease carbon dioxide emissions of road traffic—Publication of the working group of CO₂ emissions, in Finnish, Abstract in English.]
- [13] D.A. Ford, Shang inquiry as an alternative to Delphi: Some experimental findings, *Technol. Forecast. Soc. Change* 7 (1975) 139–164.
- [14] W.L. Brockhaus, A quantitative analytical methodology for judgmental and policy decisions, *Technol. Forecast. Soc. Change* 7 (1975) 127–137.
- [15] H. Sackman, *Delphi Critique. Expert Opinion, Forecasting, and Group Process*, The Rand Corporation, Lexington Books, Lexington, MA, 1975 (142 pp.).
- [16] K.Q. Hill, J. Fowles, The methodological worth of the Delphi forecasting technique, *Technol. Forecast. Soc. Change* 7 (1975) 179–192.

- [17] M. Turoff, The policy Delphi, in: H.A. Linstone, M. Turoff (Eds.), *The Delphi Method. Techniques and Applications*, Addison-Wesley Publishing, Don Mills, 1975, pp. 84–101.
- [18] H.A. Linstone, Eight basic pitfalls: A checklist in the Delphi method, in: H.A. Linstone, M. Turoff (Eds.), *Techniques and Applications*, Addison-Wesley Publishing, Don Mills, 1975, pp. 573–586.
- [19] M. Benarie, Delphi- and Delphi-like approaches with special regard to environmental standard setting, *Technol. Forecast. Soc. Change* 33 (1988) 149–158.
- [20] F. Woundenberg, An evaluation of Delphi, *Technol. Forecast. Soc. Change* 40 (1991) 131–150.
- [21] M.R. Kastein, M. Jacobs, R. van der Hell, K. Luttik, F.W.M.M. Touw-Otten, Delphi, the issue of reliability: A qualitative Delphi study in primary health care in the Netherlands, *Technol. Forecast. Soc. Change* 44 (1993) 315–323.
- [22] G. Rowe, G. Wright, F. Bolger, Delphi: A reevaluation of research and theory, *Technol. Forecast. Soc. Change* 39 (1991) 235–251.
- [23] T. Webler, D. Levine, H. Rakel, O. Renn, A novel approach to reducing uncertainty: The group Delphi, *Technol. Forecast. Soc. Change* 39 (1991) 253–263.
- [24] U.G. Gupta, R.E. Clarke, Theory and applications of the Delphi technique: A bibliography (1975–1994), *Technol. Forecast. Soc. Change* 53 (1996) 185–211.
- [25] K. Cuhls, Reasons for a new foresight approach—The FUTUR process in Germany, Paper presented at the conference *The Quest for the Futures: a Methodology Seminar*, Organised by the World Futures Studies Federation (WFSF) and Finland Futures Research Centre, Turku, Finland, June, 2000, 18 pp.
- [26] F. Karmasin, H. Karmasin, Delphi-Studie “Zukunft der Mobilität,” Bericht, Institut für Motivforschung (ifm), Archiv-Nummer: 5249, available at http://www.oeamtc.at/aka_publ, 1999, 74 pp. [The website includes also English version but it was not downloadable when the manuscript was written.]
- [27] M. Turoff, S.R. Hiltz, Computer-based Delphi processes, in: M. Adler, E. Ziglio (Eds.), *Gazing into the Oracle. The Delphi Method and its Application to Social Policy and Public Health*, Jessica Kingsley Publishers, London, 1996, pp. 56–85.
- [28] B. Schwarz, U. Svedin, B. Wittrock, *Methods in Futures Studies. Problems and Applications*, Westview Press, Boulder, CO, 1982 (175 pp.).
- [29] J.F. Preble, Public sector use of the Delphi technique, *Technol. Forecast. Soc. Change* 23 (1983) 75–88.
- [30] O. Kuusi, Expertise in the future use of generic technologies. Epistemic and methodological Considerations concerning Delphi studies, *Acta Universitatis Oeconomicae Helsingiensis A-159*, Doctorate Dissertation, Helsinki School of Economics and Business Administration, HeSE Print, 1999, 268 pp.
- [31] K. Blindt, K. Cuhls, H. Grupp, Current foresight activities in central Europe, *Technol. Forecast. Soc. Change* 60 (1999) 15–35.
- [32] M. Wilenius, J. Tirkkonen, Climate in the making: Using Delphi for Finnish climate policy, *Futures* 29 (9) (1997) 845–862.
- [33] M. Adler, E. Ziglio (Eds.), *Gazing into the Oracle. The Delphi Method and its Application to Social Policy and Public Health*, Jessica Kingsley Publishers, London, 1996 (252 pp.).
- [34] P.G. Goldschmidt, Scientific inquiry or political critique? Remarks on Delphi assessment, expert opinion, forecasting and group process by H. Sackman, *Technol. Forecast. Soc. Change* 7 (1975) 195–213.
- [35] S.D. Scheele, Consumerism comes to Delphi: Comments on Delphi assessment, expert opinion, forecasting, and group process by H. Sackman, *Technol. Forecast. Soc. Change* 7 (1975) 215–219.
- [36] R.K. Merton, P. Kendall, The focused interview, *Am. J. Sociol.* 51 (1946) 541–557.
- [37] R.K. Merton, The focussed interview and focus groups. Continuities and discontinuities, *Public Opin. Q.* 51 (1987) 550–566.
- [38] F.H. van Eemeren, R. Grootendorst, F. Snoeck Henkemans, *Fundamentals of Argumentation Theory. A Handbook of Historical Backgrounds and Contemporary Developments*, Lawrence Erlbaum Associates Publishers, Mahwah, NJ, 1996 (424 pp.).
- [39] C.H. Gladwin, *Ethnographic decision tree modeling*, Qualitative Research Methods Series 19, Sage Publications, Newbury Park, 1989 (96 pp.).

- [40] S. Hirsjärvi, H. Hurme, Tutkimushaastattelu. Teemahaastattelun teoria ja käytäntö, Helsinki University Press, 2000, 213 pp. [Research Interview. The Theory and Practise of Thematic Interview, in Finnish.]
- [41] W. Bell, Foundations of Futures Studies: Human Science for a New Era vol. 1, Transaction Publishers, New Brunswick, 1997 (365 pp.).
- [42] D.S. Scheele, Reality construction as a product of Delphi interaction, in: H.A. Linstone, M. Turoff (Eds.), The Delphi Method. Techniques and Applications, Addison-Wesley Publishing, Don Mills, 1975, pp. 37–71.
- [43] I.I. Mitroff, M. Turoff, Philosophical and methodological foundations of Delphi, in: H.A. Linstone, M. Turoff (Eds.), The Delphi Method. Techniques and Applications, Addison-Wesley Publishing, Don Mills, 1975, pp. 17–36.
- [44] J. Habermas, Theorie des Kommunikativen Handelns. Band 2. Zur Kritik der funktionalistischen Vernunft, Suhrkamp, Frankfurt am Main, 1981, 633 pp.
- [45] J. Forester, Critical Theory, Public Policy, and Planning Practise. Toward a Critical Pragmatism, State University of New York Press, Albany, 1993 (214 pp.).
- [46] R. Dubes, A.K. Jain, Validity studies in clustering methodologies, Pattern Recognit. 11 (1979) 235–254.
- [47] G. Milligan, Cluster analysis, Encyclopedia of Statistical Sciences, Wiley, New York, 1998, pp. 120–125 (Update Volume 2).
- [48] G.W. Milligan, Clustering validation: Results and implications for applied analyses, in: P. Arabie, L.J. De Soete, G. De Soete (Eds.), Clustering and Classification, World Scientific Publ., River Edge, NJ, 1996, pp. 341–375.
- [49] G. Henrion, A. Henrion, R. Henrion, Beisbiele zur Datenanalyse mit BASIC-Programmen, VEB Deutscher Verlag der Wissenschaften, Berlin, 1988 (363 pp.).
- [50] SPSS Base 7.5 Applications Guide, SPSS, Chicago, 1997 (339 pp.).
- [51] K.R. Popper, Logik der Forschung, 3rd German ed., J.C.B. Mohr (Paul Siebeck), Tübingen, 1969, 441 pp.
- [52] J. Dewey, The quest for certainty. A study of the relation of knowledge and action, Gaudeamus, Tampere, 1999 (orig. 1929; 272 pp. [Finnish Translation.]).
- [53] B. Everitt, Cluster analysis, in: D. McKay, N. Schofield, P. Whiteley (Eds.), Data Analysis and the Social Sciences, Frances Pinter, Exeter, 1983, pp. 226–255.
- [54] N. Nakicenovic, A. Grübler, A. McDonald (Eds.), Global Energy Perspectives, International Institute for Applied Systems Analysis (IIASA) and World Energy Council (WEC), Cambridge Univ. Press, Cambridge, 1998 (299 pp.).
- [55] Ministry of Transport and Communications: Liikenneskenaariot 2025, Publications of the Ministry of Transport and Communications 25/2000, Edita, Helsinki, 2000, 101 p. [Traffic Scenarios 2025, in Finnish, Summary and Conclusions in English.]
- [56] T. Meristö, Skenaariotyöskentely yrityksen johtamisessa, Acta Futura Fennica No. 3, Finnish Society for Futures Studies and VAPK-kustannus, Helsinki, 1991, 188 pp. [Scenarios in Strategic Management, Summary in English, doctorate dissertation.]
- [57] P. Tapio, Tulevaisuuden ennustamista vai tulevaisuuden tekemistä? Ympäristöongelmien haasteet tielaitoksen tulevaisuudentutkimukselle liikenne- ja ympäristöpolitiikan näkökulmasta, FinnRA Reports 63/1992, Finnish Road Administration, Helsinki, 1992, 71 pp. [Challenges that environmental problems pose to the road administration's futures research from traffic and environmental policy points of view, in Finnish, Abstract in English.]
- [58] Framtida transporter i Sverige, Väg-och trafik-Institutet, VTI notat Nr T 117-1992, 1992, 121 pp+ 2 appendices 34 pp. [Future transport in Sweden, in Swedish.]
- [59] P. Tapio, Scenarios of the traffic CO₂ policy in Finland for 2025. A paper for the Second EFIEA Climate Policy Workshop “From Kyoto to the Hague—European Perspectives on Making the Kyoto Protocol Work,” April 18–19, 2000, Amsterdam. Available at http://www.vu.nl/english/o_o/instituten/IVM/research/efiea/tapio.pdf.
- [60] P. Tapio, Climate and traffic: Prospects for Finland, Global environmental change: Human and policy dimension, forthcoming (1), (2002).